CLAIMS

I claim:

1. A method to determine the genetic diversity of a population comprising the steps:

a) providing a sample of interest containing a population of interest;

b) isolating genomic DNA from the sample;

c) performing PCR on the genomic DNA to amplify a subregion comprising a polymorphic site of interest using an upstream primer comprising a binding site for a solid support and a downstream primer to produce an amplified fragment;

d) cleaving the amplified fragment with a restriction enzyme that cleaves the amplified fragment at a site close to the polymorphic region;

e) binding the cleaved amplified fragment to the solid support to form an immobilized product;

f) splitting the immobilized product into two pools;

g) attaching a double-stranded linker to the immobilized product of each pool, wherein the linker comprises a Type IIS restriction enzyme site;

h) digesting the product with the Type IIS restriction enzyme to form a tag;

i) ligating the tags to one another to form ditags;

j) amplifying the ditags by PCR;

k) cleaving the ditags with the restriction enzyme of step c) and ligating the ditags to form a concatemer;

l) cloning the concatemer into a vector; and

m) determining the sequence the concatemer.

2. The method according to claim 1, further comprising the step of digesting the tags generated in step g) with T4 DNA polymerase to create blunt ends.

3. The method according to claim 1, further comprising the step of amplifying the ditags a second time after step j).

4. The method according to claim 1, wherein the population of interest is selected from the group consisting of a microbial population, a viral population and an immune cell population.

5. The method according to claim 4, wherein the microbial population is from a sample associated with a site for petroleum or natural gas exploration.

6. The method according to claim 4, wherein the microbial population is from a sample associated with a site of oil or gas reserves.

7. The method according to claim 4, wherein the microbial population is from a sample associated with a site for mineral exploration.

8. The method according to claim 4, wherein the microbial population is from a sample associated with an agricultural field.

9. The method according to claim 4, wherein the microbial population is from a patient sample.

10. The method according to claim 9, wherein the patient is suspected to have a bacterial or fungal infection.

11. The method according to claim 4, wherein the microbial population is from a sample associated with a bioremediation site.

12. The method according to claim 4, wherein the immune cell population is from a patient sample.

13. The method according to claim 12, wherein the patient is suffering from a disease selected from the group consisting of autoimmune disease, graft v. host disease, lymphoma, leukemia, Hodgkin's disease, cancer, atherosclerosis and psoriasis.

14. The method according to claim 4, wherein the microbial population is from a sample of an insect or a parasite.

15. A method for analyzing a marker diversity profile (MDP) by comparison to a marker diversity profile matrix database to identify at least one parameter associated with the MDP, comprising the steps of:

a) constructing a marker diversity profile comprising a plurality of nucleic acid markers,

wherein each marker is an identifier for one member or a limited number of members of a population,

wherein the abundance of said marker provides an indication as to the abundance of the corresponding member of said population, and

wherein said population is associated with at least one parameter;

b) detecting the abundance of each marker;

c) transducing the abundance of each marker into an electrical signal;

d) storing a marker diversity profile comprising the marker electrical signals in a marker diversity profile matrix data structure;

e) determining at least one parameter associated with said marker diversity profile by comparing said marker diversity profile with a

marker diversity profile matrix data structure comprising marker diversity profiles that are each known to be associated with at least one parameter.

16. The method according to claim 15, wherein the population is a microbial population.

17. The method according to claim 15, wherein a marker diversity profile is compared to a marker diversity profile matrix data structure according to comparison rules contained in a knowledge base.

18. The method according to claim 15, wherein the abundance of markers in said marker diversity profile is expressed as the relative abundance in the population.

19. The method according to claim 15, wherein the markers of said marker diversity profile are detected through DNA sequencing.

20. The method according to claim 15, wherein the markers of said marker diversity profile are detected by determining the masses of the corresponding translation products.

21. The method according to claim 15, wherein the markers of said marker diversity profile are detected through hybridization.

22. The method according to claim 15, wherein the markers of said marker diversity profile are detected through size differences of the markers.

23. The method according to claim 22, wherein the size differences of the markers correlates to a size difference of a 16S and 23S rDNA intergenic region.

24. The method according to claim 22, wherein the size differences of the markers correlates to a size difference of a 23S and 5S rDNA intergenic region.

25. The method according to claim 22, wherein the size differences of the markers are identified by denaturing gradient gel electrophoresis (DGGE) or terminal restriction fragment length polymorphism (T-RFLP).

26. The method according to claim 15, wherein said population is comprised of immunoglobulin genes.

27. The method according to claim 15, wherein said population is comprised of viral, viroid or bacteriophage genes.

28. A method for analyzing a marker diversity profile database to identify one or more robust indicator markers for a desired parameter, comprising the steps of:

a) constructing a marker diversity profile comprising a plurality of nucleic acid markers,

wherein each marker is an identifier for one member or a limited number of members of a population,

wherein the presence or abundance of said marker provides an indication as to the presence or abundance of the corresponding member of said population, and

wherein said population is associated with said desired parameter;

b) detecting the presence or abundance of each marker;

c) transducing the presence or abundance of each marker into an electrical signal;

d) storing the marker diversity profile comprising marker electrical signals in a marker diversity profile matrix data structure;

e) correlating the presence or abundance of each marker in the marker diversity profile database with the presence or abundance of each said parameters;

f) ranking the markers for each parameter based upon correlation values; and

g) designating one or more markers as robust indicators for the parameter as those that have a correlation value of at least 0.4 to the parameter.

29. The method according to claim 28, wherein the correlation value is at least 0.5.

30. The method according to claim 28, wherein the correlation value is at least 0.6.

31. The method according to claim 28, wherein the correlation value is at least 0.7.

32. The method according to claim 28, wherein the correlation value is at least 0.8.

33. The method according to claim 28, wherein the correlation value is at least 0.9.

34. The method according to claim 28, wherein the correlation value is at least 0.95.

35. The method according to claim 28, wherein the presence of said marker correlates with either the presence or the abundance of said parameter.

36. The method according to claim 28, wherein the abundance of said marker correlates with either the presence or abundance of said parameter.

37. A method for analyzing a MDP database to identify groups of correlated markers, comprising the steps of:

a) constructing a marker diversity profile comprising a plurality of nucleic acid markers,

wherein each marker is an identifier for one member or a limited number of members of a population,

wherein the presence or abundance of said marker provides an indication as to the presence or abundance of a corresponding member of said group, and

wherein said group is associated with a desired parameter;

b) detecting the presence or abundance of each marker;

c) transducing the presence or abundance of each marker into an electrical signal;

d) storing the marker diversity profile comprising marker electrical signals in a marker diversity profile matrix data structure;

e) correlating the presence or abundance of each marker in the marker diversity profile database with the presence or abundance of other markers in pairwise comparison to form a correlation matrix;

f) separating the markers into correlated groups based upon values from correlation matrix.

38. A method of constructing a matrix diversity profile database, comprising the steps of

a) constructing a marker diversity profile comprising a plurality of nucleic acid markers,

wherein each marker is an identifier for one member or a limited number of members of a population,

wherein the presence or abundance of said marker provides an indication as to the presence or abundance of the corresponding member of said population, and

wherein said population is associated with at least one parameter;

b) detecting the presence or abundance of each marker;

c) transducing the presence or abundance of each marker into an electrical signal;

d) storing the marker diversity profile comprising marker electrical signals in a marker diversity profile matrix data structure; and

e) indexing the marker diversity profile with the associated parameter to form a matrix diversity profile database.

39. A method for identifying a marker that correlates with a desired parameter, comprising the steps of

a) constructing a marker diversity profile comprising a plurality of nucleic acid markers,

wherein each marker is an identifier for one member or a limited number of members of a population,

wherein the presence or abundance of said marker provides an indication as to the presence or abundance of the corresponding member of said population, and

wherein said population is associated with said desired parameter;

b) detecting the presence or abundance of each marker;

c) correlating the presence or abundance of each marker in the marker diversity profile database with the parameter;

d) ranking the markers for the parameter based upon correlation values; and

e) designating one or more markers as a robust indicator for the parameter as those that have a correlation value of at least 0.4 to the parameter.

40. The method according to claim 39, wherein said desired parameter is the presence of mineral deposits and/or petroleum reserves.

41. A method for identifying a marker that correlates with a desired parameter, comprising the steps of

a) constructing a marker diversity profile comprising a plurality of nucleic acid markers,

wherein each marker is an identifier for one member or a limited number of members of a population,

wherein the presence or abundance of said marker provides an indication as to the presence or abundance of the corresponding member of said population, and

wherein said population is associated with said desired parameter;

b) detecting the presence or abundance of each marker;

c) correlating the presence or abundance of each marker in the marker diversity profile database with the parameter;

d) ranking the markers for the parameter based upon correlation values;

e) designating one or more markers as a potential indicator for the parameter as those that have a correlation value of at least 0.4 to the parameter; and

f) determining whether the potential indicator is present or abundant in a population that is not associated with said desired parameter; and

g) designating one or more markers as a robust indicator for the parameter if the potential indicator is not present or is less abundant in a population that is not associated with said desired parameter than in a population that is associated with said desired parameter.

42. The method according to claim 41, wherein said desired parameter is the presence of mineral deposits and/or petroleum reserves.

43. A method for performing mineral deposit exploration or subsurface petroleum or natural gas reserve exploration, comprising the steps of:

a) isolating nucleic acids from an environmental sample;

b) amplifying a specific DNA molecule that has been shown to correlate with the presence of mineral deposits and/or petroleum or natural gas reserves;

c) detecting the presence of said amplified DNA molecule; and

c) inferring the presence of said mineral deposits and/or said petroleum or natural gas reserves from the presence of said amplified DNA molecule.

44. A method for performing mineral deposit exploration or subsurface petroleum or natural gas reserve exploration, comprising the steps of:

a) isolating nucleic acids from an environmental sample;

b) amplifying a specific RNA molecule that has been shown to correlate with the presence of mineral deposits and/or petroleum or natural gas reserves;

c) detecting the presence of said amplified RNA molecule; and

c) inferring the presence of said mineral deposits and/or said petroleum or natural gas reserves from the presence of said amplified RNA molecule.